

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2001, 00:19:43 ; Search time 1440.11 Seconds

(without alignments)
10309.950 Million cell updates/sec

Title: US-09-510-332-2

Perfect score: 900
1 atgcataagctcaccatc.....acagtaagctcgtcagta 900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sta: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_ov: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sta: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_hgo_hum: *
31: em_hgo_inv: *
32: em_hgo_inv: *
33: em_hcg_hum: *
34: em_hcg_inv: *
35: em_hcg_inv: *
36: em_hcg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	900	100.0	900	6	AX097696	AX097696 Sequence
2	900	100.0	900	6	AF227129	AF227129 Homo sapi
3	900	100.0	132953	9	AC003015	AC003015 Human BAC
4	898.4	99.8	146296	9	AC026787	AC026787 Homo sapi
5	675.6	75.1	141747	2	AC034214	AC034214 Homo sapi
6	390.2	43.4	1008	10	AF227140	AF227140 Homo sapi
7	390.2	43.4	1331	6	AX097772	AX097772 Sequence
8	388.6	43.2	1008	10	AF227149	AF227149 Mus muscu
9	388.6	43.2	1698	6	AX097836	AX097836 Sequence
10	124.4	13.8	957	6	AX097708	AX097708 Sequence
11	124.4	13.8	957	9	AF227133	AF227133 Homo sapi
12	124.4	13.8	173735	9	AC006518	AC006518 Homo sapi
13	124.4	13.8	190002	2	AC068649	AC068649 Homo sapi
14	123.2	13.7	939	6	AX097712	AX097712 Sequence
15	123.2	13.7	939	9	AF227135	AF227135 Homo sapi
16	119.8	13.3	930	6	AX097751	AX097751 Sequence
17	119.8	13.3	951	6	AX097700	AX097700 Sequence
18	119.8	13.3	951	9	AF227130	AF227130 Homo sapi
19	119.8	13.3	82419	9	AC004979	AC004979 Homo sapi
20	117.8	13.1	167869	9	AC073342	AC073342 Homo sapi
21	117.8	13.1	184034	2	AC024156	AC024156 Homo sapi
22	111.4	12.4	213278	2	AC018630	AC018630 Homo sapi
23	108.8	12.1	3775	6	AX097857	AX097857 Sequence
24	108.8	12.1	960	6	AX097746	AX097746 Sequence
25	107.8	12.0	179029	2	AC079625	AC079625 Homo sapi
26	107.2	11.9	903	6	AX097723	AX097723 Sequence
27	106.8	11.9	1103	6	AX097848	AX097848 Sequence
28	103.6	11.5	2567	6	AX097840	AX097840 Sequence
29	102	11.3	184034	2	AC024156	AC024156 Homo sapi
30	101.4	11.3	1316	6	AX097830	AX097830 Sequence
31	101.2	11.2	894	10	AF240766	AF240766 Rattus no
32	101.2	11.2	1304	6	AX097782	AX097782 Sequence
33	101	11.2	314700	2	AC018743	AC018743 Homo sapi
34	95.4	10.6	162380	2	AC018851	AC018851 Homo sapi
35	95.4	10.6	314700	2	AC018743	AC018743 Homo sapi
36	94.2	10.5	912	6	AX097719	AX097719 Sequence
37	94.2	10.5	912	9	AF227137	AF227137 Homo sapi
38	93.8	10.4	162380	2	AC018851	AC018851 Homo sapi
39	92.4	10.3	900	10	AF227141	AF227141 Rattus no
40	92.4	10.3	6552	6	AX097776	AX097776 Sequence
41	91.4	10.2	900	6	AX097704	AX097704 Sequence
42	91.4	10.2	900	9	AF227132	AF227132 Homo sapi
43	90.6	10.1	900	6	AX097748	AX097748 Sequence
44	90.6	10.1	930	6	AX097757	AX097757 Sequence
45	87.8	9.8	924	6	AX097714	AX097714 Sequence

ALIGNMENTS

RESULT 1

AX097696 LOCUS AX097696 900 bp DNA 30-MAR-2001

DEFINITION Sequence 2 from Patent WO0118050.

AX097696 ACCESSION AX097696 GI:13514345

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

1.900

/organism="Homo sapiens"

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Zuker, C.S., Adler, J.E., Ryba, N., Mueller, K. and Hoon, M.
T2R taste receptor family
Patent: WO 0118050-A 2 15-MAR-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (US) ; THE SECRETARY OF
THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
Location/Qualifiers

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	301	GTTCCTCTTTGTCGCAAGTGTGCAGCGTCGTCACCCACTTTCATCTGATGTTGAAGATG	360													
Qy	361	aggatataccaagccggtcccaatggaatgatcctggtgctctgcatatgataatcatt	420													
Db	361	AGGATATCCAAAGCGTGTCCTCATGATGATCCCTGGGGTCTGCTATATGATCATGATG	420													
Qy	421	tgatcttcccaatgcaaatgacaggttttcttgctccatctctccaaggaatttttc	480													
Db	421	TGATCTTTCATACCAAAATATGACAGGGTTTATGGTCCATATCTTCATAGGAAATTTTTC	480													
Qy	481	tcccaaatgccaacatccaataaagaatatcactggtgatacagattctctctgt	540													
Db	481	TCCCAAAATGCCACAAATTCAAAAAGAAATACACTGGCTATACAGATTTTCTTTTGT	540													
Qy	541	gcctgaatctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc	600													
Db	541	GCTGAGTCTCAGTGCCATGCTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	600													
Qy	601	ctggggagagcaacccggcaaatgagaacaacagtgccggcagcagaggtcttcgca	660													
Db	601	CTGGGAGAGCACACCCGGCAAAATGAGAAACACAGTGCCGGCAGCAGAGGTTCTGGCAG	660													
Qy	661	ggtgaccccatcagcaggtgctgctatcctgctcctccgactcctcattctccac	720													
Db	661	GGTGACCCCATCAGCGGCTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720													
Qy	721	tgcataataaagtttct	780													
Db	721	TGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	780													
Qy	781	tctctacatctctgattggtatataacccttcggaacctctcattcatttaattaga	840													
Db	781	TTCCTCATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	840													
Qy	841	aatcctaataatgaaacaaatgcaaaaagtctcctccacagtaagtgttcagaga	900													
Db	841	AATCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	900													
RESULT	3															
LOCUS	AC003015	132953 bp	DNA													
DEFINITION	Human BAC clone GSI-113H23 from 5p15.2, complete sequence.															
ACCESSION	AC003015															
VERSION	AC003015.1	GI:2547255														
KEYWORDS	HTG.															
SOURCE	human.															
ORGANISM	Homo sapiens															
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;															
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.															
TITLE	1 (bases 1 to 132953)															
JOURNAL	Wamsley, P., Kramer, J., Elliott, G. and O'Brien, D.															
REFERENCE	The sequence of H. sapiens BAC clone GSI-113H23															
AUTHORS	Unpublished															
TITLE	2 (bases 1 to 132953)															
JOURNAL	Waterston, R.															
REFERENCE	Direct Submission															
AUTHORS	Submitted (20-OCT-1997) Department of Genetics, Washington															
TITLE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA															
JOURNAL	3 (bases 1 to 132953)															

```

Web site: http://genome.wustl.edu/gsc
Contact: saplens@watsn.wustl.edu
----- Summary Statistics -----
Center project name: H_GS113H23
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone; and the assembly was
confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University School of
Medicine, St. Louis MO.

Mapping information for this clone was also provided by Dr. Michael
Lovett, Departments of Otorhinolaryngology, Molecular Biology and
Oncology, University of Texas Southwestern Medical Center, Dallas
TX.

SOURCE INFORMATION:
This clone is from the first BAC library from Genome Systems, Inc.
(http://www.genomesystems.com).
Cell line: Lymphoblastoid
Haplotypes: two
VECTOR: pBelobAC
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is GS1-330J10, 200 bp overlap.
Actual start of this clone is at base position 1 of GS1-113H23;
actual end is at 132953 of GS1-113H23.

This clone contains SMS HSC022YA5 (MID:g125481) and HS268ZD9
(MID:g454585).

This clone contains polymorphisms with GS1-330J10.
Location/Qualifiers
1. 132953
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="GS1-113H23"
/clone_lib="GSBAC1"
/map="5p15.2"
complement(20..194)
/rpt_family="L1"
195..358
/rpt_family="L1"
/rpt_family="ALU"
complement(359..853)
/rpt_family="L1"
2903..2933
/rpt_family="L1"
complement(3640..3667)
/rpt_family="L1"
complement(3764..3788)
/rpt_family="L1"
complement(3820..4112)
/rpt_family="ALU"
complement(4332..4351)
/rpt_family="L1"
complement(4444..4464)
/rpt_family="L1"
complement(5046..5122)
/rpt_family="L1"

```


Db	49767	ATTTCACAAATAGGCATCATTTGGTGGGAATGGCATTTGACTTGAATACAGCACAGAAAA	49708
Qy	121	atggtccggctggatctcccttctcttgctcggagttctagaattttctcgagttg	180
Db	49707	ATGGCTCCGGCTGGATCTCCTTTCTTGTCGGCAGTTTCTAGAAATTTTCTGCAGTTG	49648
Qy	181	ttcatctcttaagtaagttgtgttctctctcatagaattcaatgtgttcg	240
Db	49647	TTCAATCTTCTACGTTAAATGTGATTTGTATCTTTCATAGAAATTCATATGATGTTCTCG	49588
Qy	241	aattgtgcaattctctctttaaaatgaatggaacttggctcgcacaaatgctgcg	300
Db	49587	AATGTGCATATCTCTTATTTATTAATGATTAATGGAACTTTGGCTTGCCACATAGCTCGGC	49528
Qy	301	gtttctatttggccaaagtttgcgcggtcgtccacacatcttcaatctggttgaagtg	360
Db	49527	GTTTTTCATATGGCCAAAGTGTGCCAGCGCTCCATCCACCACTTCATCTGTGGTGAAGATG	49468
Qy	361	aggatatacacaagctgtgccaatgatagtacctggggtctctgctaatagtatctg	420
Db	49467	AGGATATCCAAAGCTGTGTCCATGATGATCCTGGGGTCTTGCTATATGATCTATGATTT	49408
Qy	421	tgtgtttccatagcaaatatgcaaggtttaaitygtccatacttctcaaggaaatttctc	480
Db	49407	TGTGTTTTCCATAGCAAAATATGCAGGGTTTATGCGCCATACTCTCTAAGAAATTTTTC	49348
Qy	481	tcccacaaatgccaacatctcaaaaagaagatacactggtatacagaatttctcttgct	540
Db	49347	TGCCAAATATGCCACATTTCAAAAAGATATACACTGGCTATACAGATTTTCTCTTTGTT	49288
Qy	541	gctgaagctctcaagtgccattgcttactctctcttcttgcttgcttgccttgatctct	600
Db	49287	GCTGAGTTCTCAGTGCATGSCCATGTGCTATCTTCCTTTTGGTGTGCTTGATTTTCTCT	49228
Qy	601	ctgggggaagcaccacccggaatgagaacacaaatgctggcgcagcagaaggtctctgcagg	660
Db	49227	CTGGGGAAGGCACACCCGGCAATGAGAAACACAGAGGCGGACAGCAGGATTCCTGGCAGG	49168
Qy	661	ggtgaaccatcagcaggttgctgtctactctgtccctccgatccctactctctccac	720
Db	49167	GGTGACACCCATCAGGCGGTGCTGTCTATCCGTCGTCCTTCGATGCTCTACTTCTCCAC	49108
Qy	721	tgcatagataaaaagtttctctctctctctcttaaaagtttcaatcagaagaagttcaatcttg	780
Db	49107	TGCATGATAAAAAGTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG	49048
Qy	781	ttctcatcctctgtgattggtgataaccctcttgaaactcttcaatcttaatttga	840
Db	49047	TTCTTCATCCTTGTGATTTGGATATACCTCTCTGACACTCTTCATCTTAAATTTTATGA	48988
Qy	841	aatcctaaattgaaacaaatgycacaaaagttctctctccacaglaagtgtctgcagtga	900
Db	48987	AATCCTAAATTTGAAACAAATATGCAAAAAAGTTCTCTCTCCACAGTAAAGTGTGCAGTGA	48928
RESULT 5			
AC034214			
LOCUS			
DEFINITION			
Homo sapiens chromosome 5 clone CTD-2001E22, WORKING DRAFT			
SEQUENCE, 16 ordered pieces.			
AC034214			
AC034214			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			

Dd	421	ACTTTCATCCATAGCAGAGACCTTCAGCAATCCTTAAACCATTATTATAAGCCTTTT	480
Oy	481	lcccaaatgcacaattccaagaagatatacactggtctatacagatttctcttggt	540
Dd	481		
Dd	481	CCTAATAAATG---CACTCATAGTCGGAAACAGGGCATGCACACTACTCATGCTCGTGC	537
Oy	541	gctgaagtctcaagtcacatlgcttatcttctcttttgtcttgcttgccttgattctc	600
Dd	538	CTTGCGGTACACACTGCCGTTGTTATCTTTACTGTTGCTGTTCTGCTTGATTAATCTCC	597
Oy	601	ctggaggagcacaccggccaatgagaacacacagtgcccgagcagcagggttccctgcaag	660
Dd	538		
Dd	538	CTGTGAATTTATAGCAGCACAGATGAGGACTA---TGGTAGGCACACAGGAGATATGCGGA	654
Oy	661	ggtgcaccatcaagcgttgctgctaactcgtctctccctccgatcccttaactccccac	720
Dd	655		
Dd	655	CATGCTCATATCAGTGCAGATGCTGTCATTTATCATCTTCATCTATCTCTCCAC	714
Oy	721	tgcataataaaagtttctctctctctcttcttaaagtttccatcagaaggttcatcttcg	780
Dd	715		
Dd	715	TACATGGTGGCTGTTCTCATCTCATCTCAAGTCCCTCTACTTGGACGACAAGCACTTTGTA	774
Oy	781	ttcttcacaccttgatggtgatataccctctctgacacctcttcataattttaaga	840
Dd	775		
Dd	775	TTCTCTTACTAGTGTATTGTGATATGATACCCCTCAATACATCGATTGTCTTAATTTAGA	834
Oy	841	aatcctaaattgaaaacaaatgacaaaaaagttcctctcccacagtaagtgctgcagty	899
Dd	835		
Dd	835	AATCCTAAGCTGAACGAAATGCAAAAATGTTCAATGTCATGTAAGTGTGTCATG 893	
RESULT	7		
LOCUS	AX097772	1331 bp	DNA PAT 30-MAR-2001
DEFINITION	Sequence 78 from Patent WO0118050.		
ACCESSION	AX097772		
VERSION	AX097772.1	GI:13514400	
KEYWORDS			
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
	Eularypota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 1331)		
AUTHORS	Zuker,C.S., Adler,J.E., Ryba,N., Mueller,K. and Hoon,M.		
TITLE	T2r taste receptor family		
JOURNAL	Patent: WO 0118050-A-78 13-MAR-2001.		
	THE REGENTS OF THE UNIVERSITY OF CALIFORNIA (US) ; THE SECRETARY OF		
	THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
FEATURES	Location/Organism		
source	1..1331 /organism="Rattus sp." /db_xref="taxon:10118" /note="rat T2R01, rGR01"		
BASE COUNT	348 a 289 c 266 g 428 t		
ORIGIN			
Query Match	43.4% Score 390.2; DB 6; Length 1331;		
Best Local Similarity	66.2%; Pred. No. 5.7e-83;		
Matches 595; Conservative 0; Mismatches 256; Indels 6; Gaps 2;			
Oy	1 atgctagagctcaccctatcatctatttctctctgcagtgatacaattctcttgagg	60	
Dd	77 ATGATGAGAGGAGCATATACCTCTCTTTTGGTTGGATGGATGGACATTTGTCACTGGG	136	
Oy	61 atttccaaatgagcatcatgtgtgtgtgaatgagcatgtgactgataagcagaanaa	120	
Dd	137 GTCTTGGCAAATGAGCCTCATATGTGCTGTGCATCTAATGACTTGATCATGTGGAAGA	196	
Oy	121 atggctccgctgagatctctctctctctctctgctgcagttcttagaatcttctgcagty	180	


```

SOURCE          1. 1096
                /organism="Mus sp."
                /db_xref="taxon:1095"
                /note="mouse T2R19, mGRI9"
BASE COUNT      426 a      397 c      352 g      523 t
ORIGIN

```

QY 1 atgcagagctcaccatcatatcatcttcttcgtgacgigalacaaattctcttg 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db . 515 ATGATGAAGCTCATATGCTCTTCCCTTCGTGGTGCGTAGTCAGCTTTTAAC TGG 574

Db 875 AGGATATCCAGGTTGCTGCCATGGCTGATCCCTGGCATCTGTGCTCTATGTAAGTCTACT 934

[illegible]

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="human T2R07, hGR07"
BASE COUNT      258 a      202 c      192 g      305 t
ORIGIN

```

[illegible]

Db 137526 TTTTTCCTTGATGCTCTCCCTGGGAGACATATACAGGCAATGACGTCAGTCCACA 137585
 Oy 640 ggcagcagggttcctcgcagagggtgacacacacacgcgcgtcgtctatcctcctc 699
 Db 137586 GCGTGCAGAGAGCCCGACGAGAGCCCATCTGAGAGCCCGAAGAGCTCATTTTCCTTC 137645
 Oy 700 ctgtcctccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 759
 Db 137646 CTTCCTCTCTTATTTGCTTACTATTTGCTTCTTCTCATTTGCCACTCCAGCTATTATG 137705
 Oy 760 atcgaaggttcacatctctcctcctcctcctcctcctcctcctcctcctcctc 816
 Db 137706 CCAGAGACGCAATTAAGCTGCTGATTTTTCGTCAGTCATGCTTAATCTACCCCAAGT 137765
 Oy 817 cactctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 872
 Db 137766 CATTCATTTATTCATTAATCTGCGGACAAATTAATTAACATGCTATCTTAAGGT 137821

RESULT 14

LOCUS AX097712 939 bp DNA PAT 30-MAR-2001
 DEFINITION Sequence 18 from Patent WO0118050.
 ACCESSION AX097712
 VERSION AX097712.1 GI:13514362
 KEYWORDS
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 939)
 ZUKER,C.S., Adler,J.E., Ryba,N., Mueller,K. and Hoon,M.
 T2R taste receptor family.
 Patent: NO 0118050-A-18 15-MAR-2001.
 TITLE THE RECEPTORS OF THE UNIVERSITY OF CALIFORNIA (US) ; THE SECRETARY OF
 JOURNAL THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

FEATURES
 SOURCE
 1..939
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="human T2R9, hGR9"

BASE COUNT 248 a 187 c 186 g 318 t
 ORIGIN

Query Match 13.7% Score 123.2; DB 6; Length 939;
 Best Local Similarity 49.2% Freq. No. 3.7e-19;
 Matches 419; Conservative 0; Mismatches 418; Indels 15; Gaps 3;
 Oy 22 atcctctctcctcctcgcagtgatcacatctcctcctcctcctcctcctcctcctc 81
 Db 22 AT 81
 Oy 82 ggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 141
 Db 82 GACTAGTAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 141
 Oy 142 cttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 193
 Db 142 CTGATCAGCTTGGCCATTCAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 201
 Oy 194 ttaagtgatgtatcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 253
 Db 202 TTTATGCTCTCTCTTCCAGAGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261
 Oy 254 tcttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 309
 Db 262 GTCTGACATTTTGCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 321
 Oy 310 tggcgaaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 369
 Db 322 TTACTCAAGATAGCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 381

Oy 370 aagctggtcccatgatgatccctcctcctcctcctcctcctcctcctcctcctcctc 429
 Db 382 AAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
 Oy 430 catagcaaatatgcaaggttctatggtcccatcctcctcctcctcctcctcctcctc 489
 Db 442 AAGAAATATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
 Oy 490 gcccaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 549
 Db 502 AAATTCAAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 561
 Oy 550 tcaagtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 609
 Db 562 ATGATGCTCTTATTCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
 Oy 610 cacacccgcgaatggaagaacacagtgccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 669
 Db 622 CACACCAAGCAGATTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
 Oy 670 atcagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 726
 Db 682 ATGAGGCCATTAAGGCGAGTATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
 Oy 727 ataaagttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 786
 Db 742 TTTCTGTTATGACCTCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
 Oy 787 atccttgatgtgatatatcctcctcctcctcctcctcctcctcctcctcctcctc 846
 Db 802 GACATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 861
 Oy 847 aatggaacaa 858
 Db 862 AAGTTGAGGAA 873

RESULT 15

LOCUS AF227135 939 bp DNA PRI 18-MAR-2000
 DEFINITION Homo sapiens candidate taste receptor T2R9 gene, complete cds.
 ACCESSION AF227135
 VERSION AF227135.1 GI:7262616
 KEYWORDS
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 939)
 Adler,E., Hoon,M.A., Mueller,K.L., Chandrasekhar,J., Ryba,N.J. and
 Zuker,C.S.
 A novel family of mammalian taste receptors
 Cell 100 (6), 693-702 (2000)

TITLE

2 (bases 1 to 939)
 Adler,E., Hoon,M.A., Mueller,K.L., Chandrasekhar,J., Ryba,N.J. P.
 Direct Submission
 Submitted (21-JAN-2000) NIDCR, NIH, Bldg. 10 Room 1N106, 10 Center
 Drive, Bethesda, MD 20892, USA

JOURNAL

Location/Qualifiers
 1..939
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="12"
 /map="12p13"

FEATURES

source
 primer_bind
 1..19
 1..939

/note="G protein-coupled receptor; PCR-derived sequence;
 similar to an interval in BAC AC006518"
 /codon_start=1
 /product="candidate taste receptor T2R9"

```

/protein_id="AAFA43908.1"
/db_xref="GI:7262617"
/translation="MPSAIEAIVYIIILAGELTTIGMNGFIVLVNCIDMLKRRDISLIT
DIILSLAISIRICLCLVSLDGFPMFLPPGYGASVLSIVYVWFAANNSLWETSCG
LSIYFLUKTIANISHPPEFMILKINKWALILCSPLISLIISVPKDDMMYHIFKXS
HEENITWKEKYSIPGTEKOLLNIGWNPVILCLISFLLBSLVYHTKOIRLHANG
FRDSTEAHMAIKAVIIFLLLIYVIVPVLVMTSSALIPGKLVLMIGDIVIYFPPS
SHSLILMGNSKLEAEFLKMLRFVCKFLRRKPPVP"
<1..>939
mRNA
/product="candidate taste receptor T2R9"
/primer_bind
BASE COUNT 248 a 187 c 186 g 318 t
ORIGIN

```

Query Match: 13.7% Score 123.2; DB 9; Length 939;
Best Local Similarity 49.2%; Pred. No. 3.7e-19;
Matches 419; Conservative 0; Mismatches 418; Indels 15; Gaps 3

QY	22	atcattctctctctgacgtgatgatacaattctctctgaggaattctccaaatgacatcatt	81
DB	22	ATATATATATATTTTAATTCCTGCTGTAATGACCAATAGGCAATTTGGGCAATGATTCAAT	81
QY	82	gtcggtgtgaatgaggaattgacttgatccaagcagaaaaatggctccgctgatactcctt	141
DB	82	GTACTATGTTAACTGCAATTGACTGCGCCAAAGAAAGATATTTCTGTATTACATCATC	141
QY	142	ctctctgtctgcgcagtttctagaaatttcttcgcagttgcatctctcaag-----	193
DB	142	CTGATCAGCTTGGCCATCTCCAGAAATCTGTGCTGTGTATATCATTAATGCTTC	201
QY	194	ttaagtgtgtgttbaatctctctctatagaatcaatcaatgctgtctgcgaattgtgcaatc	253
DB	202	TTTATGTCCTCTTTCCACAGTACATATGGCAATACGTCGTGATAGCAATTTTGATGAT	261
QY	254	tcctt-----attataaalgaaltggaactctgacctgcgcacatggctcgcccttcat	309
DB	262	GTCTGGACATTTGCCAATATTCAGATGCTCTGTTACTTCTTGCTCGATCTCTGAT	321
QY	310	tgtgcgaaggtgtgcgaaggtgcgtaccacaccttcatctggttgaagatgagatcc	369
DB	322	TTTACTCAAGATAGCCAAATATATCGACACCATTATTTCTTGCTGACTAAAGATCAAC	381
QY	370	aagctgtgtcccatgtagatcctggtggtctctgcatatgtatcatgattgtgttttc	429
DB	382	AAGGTCAAGCTTGGCATTCCTCTGGGGTCCTTTTATCTCTTAAATTATGATTGCCA	441
QY	430	catagcaaatgacgaggtttatgtgtcccatcattctccaagaaattttctcccaaat	489
DB	442	AMGATGTGATATGTGTGATATCACTTTTCAAAGTCAGTACGAAAGAAACATTTACTGG	501
QY	490	gccacaattccaanaaagaagatacaactgtgctatacgaattctctctgtgtgtagctc	549
DB	502	AAATTCAAAGTAGATMAAATTCAGAGTACTTTCAAACAGTTAACCTTAACCTGGGGGTG	561
QY	550	tcaagtgcaatgctatctctctctttgtcgtgttgcctctgattctctctcctcgggaag	609
DB	562	ATGGTTCCTTTATTCCTTTGGCTGTATCTATTTTCTTTGTTACTTTTCTCCCTAGTTAGA	621
QY	610	cacacccggcaaatgagaacacagtggtcgcgcaaggtgttcctgggaaggtgtgcacc	669
DB	622	CACACCAAGCAGATTCGACTGCATGCTCAGAGGTTCAAGAGACCACAGTACGAGGCCAC	681
QY	670	atcagcgaggtgtcgtctatcctgtctctctctgatacctctactc--tcccaatgcatg	726
DB	682	ATGAGGGCCATTAAGGCGATATCATCTTTTCTGCTCCCTCCATCATGTTACTTACCAAGTC	741
QY	727	ataaagatttctctctctctctaaagttctacatacgaaggttccatcttctgtcttc	786
DB	742	TTTCTTTGTTATGACCTCTAAGCGCTGTGATTTCTCCAGGAAATATGATGTTATGATGGT	801
QY	787	atccttgtagattgtatataccctctcggacacctctcatccttaattttaggaatcct	846

DB	802	GACATAGTAAGTGCATTTTCCCATCAAGCATTCATTCATTCATTAATATGGAATATGC	861
QY	847	aaattgaacaa	858
DB	862	AAGTTGACGGAA	873

```
Search completed: December 11, 2001, 02:40:45
Job time: 8462 sec
```

Tue Dec 11 09:54:24 2001

us-09-510-332-2.rge

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 11, 2001, 01:03:48 ; Search time 144.69 Seconds

(without alignments)
5332.730 Million cell updates/sec

Title: US-09-510-332-2

Perfect score: 900

Sequence: 1 atcctagagctcactcat.....acagtaagctcgtcagtgta 900

Scoring table:

IDENTITY_NUC
Gap 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Genes2_1101:*

- 1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	900	100.0	900	22	AAF92502	Human T2R01 nucleio
2	390.2	43.4	1331	22	AAF92528	Rat T2R01 nucleio
3	388.6	43.2	1698	22	AAF92560	Mouse T2R19 nucleio
4	124.4	13.8	957	22	AAF92508	Human T2R09 nucleio
5	123.2	13.7	939	22	AAF92510	Human T2R09 nucleio
6	119.8	13.3	930	22	AAF92525	Human T2R03 nucleio
7	119.8	13.3	951	22	AAF92506	Human GTP-binding
8	119.8	13.3	951	22	AAF92504	Human T2R03 nucleio
9	108.6	12.1	3775	22	AAF92571	Mouse T2R03 nucleio
10	108.6	12.1	960	22	AAF92523	Human T2R03 nucleio
11	108.4	12.0	2695	20	AAZ20861	Polynucleotide seq

12	107.2	11.9	903	22	AAF92515	Human T2R15 nucleio
13	106.8	11.9	1103	22	AAF92566	Mouse T2R25 nucleio
14	103.6	11.5	2567	22	AAF92562	Mouse T2R21 nucleio
15	101.4	11.3	1316	22	AAF92557	Mouse T2R16 nucleio
16	101.2	11.2	1304	22	AAF92533	Rat T2R06 nucleio
17	94.2	10.5	912	22	AAF92513	Human T2R13 nucleio
18	92.4	10.2	6552	22	AAF92530	Rat T2R03 nucleio
19	91.4	10.2	900	22	AAF92506	Human T2R05 nucleio
20	90.6	10.1	900	22	AAF92524	Human T2R31 nucleio
21	90.6	10.1	930	22	AAF92527	Human T2R37 nucleio
22	89.4	9.9	1260	22	AAO05499	Human secreted pro
23	88.6	9.8	1122	22	AAO05535	Human secreted pro
24	87.8	9.8	924	22	AAF92511	Human T2R10 nucleio
25	87.4	9.7	930	22	AAF92526	Human T2R36 nucleio
26	87.2	9.7	930	22	AAF92509	Human T2R08 nucleio
27	86.8	9.6	1354	22	AAF92558	Mouse T2R17 nucleio
28	86.4	9.6	1482	22	AAF92556	Mouse T2R15 nucleio
29	86.2	9.6	2887	22	AAF92559	Mouse T2R18 nucleio
30	85.2	9.5	1339	22	AAF92550	Mouse T2R09 nucleio
31	85	9.4	1371	22	AAF92551	Mouse T2R10 nucleio
32	84.8	9.4	954	22	AAF92514	Human T2R14 nucleio
33	84.4	9.4	936	22	AAF58252	Oligonucleotide D1
34	84.4	9.4	936	22	AAF58254	Oligonucleotide D1
35	84.4	9.4	936	22	AAF58257	Oligonucleotide D1
36	84.4	9.4	936	22	AAF58259	Oligonucleotide D2
37	84.4	9.4	936	22	AAF58262	Oligonucleotide D2
38	84.4	9.4	938	22	AAF58255	Oligonucleotide D1
39	83.8	9.3	910	22	AAF92503	Human T2R02 nucleio
40	83	9.2	1199	22	AAF92554	Mouse T2R03 nucleio
41	82.6	9.2	1759	22	AAF92543	Mouse T2R02 nucleio
42	81.6	9.1	1374	22	AAF92542	Mouse T2R01 nucleio
43	80.6	9.0	936	22	AAF58252	Oligonucleotide D1
44	80.6	9.0	936	22	AAF58254	Oligonucleotide D1
45	80.6	9.0	936	22	AAF58257	Oligonucleotide D1

ALIGNMENTS

RESULT 1	
ID	AAF92502 standard; DNA; 900 BP.
AC	AAF92502;
XX	
DT	16-MAY-2001 (first entry)
DE	Human T2R01 nucleotide sequence SEQ ID NO:2.
XX	
KW	Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor;
KW	taste transduction G-protein coupled receptor; identification; tongue;
KW	taste sensory neuron; taste cell; taste modulator; food;
KW	taste signaling pathway; ds.
OS	Homo sapiens.
XX	
PN	WQ200118050-A2.
XX	
PD	15-MAR-2001.
XX	
PF	08-SEP-2000; 2000MO-US24821.
XX	
PR	10-SEP-1999; 99US-0393634.
XX	
PR	22-FEB-2000; 2000US-0510332.
XX	
PA	(BEGC) UNIV CALIFORNIA.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Zuker CS, Adler JE, Ryba N, Mueller K, Hoon M;
DR	WPI: 2001-211396/21.
XX	
XX	F-PSDB; AAB87731.

PR Nucleic acids encoding the T2R family of G-protein coupled taste
 PR receptors, useful for identifying taste modulators that can be used in
 PR food and pharmaceutical industries to customize taste, for e.g. to
 PR decrease the bitter taste of food -
 PS Disclosure; Page 160; 249pp; English.
 XX
 CC AAF92502 to AAF92572 represent nucleic acids which encode taste
 CC transduction G-protein coupled receptors designated T2R proteins.
 CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
 CC represent T2R family consensus sequences from the present invention.
 CC The T2R proteins are taste modulators. The nucleic acids are useful as
 CC probes for the identification of taste cells, as the nucleic acids are
 CC specifically expressed in taste cells. They also serve as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centres in the brain. The taste modulators are useful for
 CC pharmacological and genetic modulation of taste signalling pathways.
 CC Modulatory compounds comprising T2R proteins can therefore be used in
 CC food and pharmaceutical industries to customise taste, for e.g. to
 CC decrease the bitter taste of food or drugs.
 XX
 SO Sequence 900 BP; 203 A; 203 C; 172 G; 322 T; 0 other;

Query: Match 100.0%; Score 900; DB 22; Length 900;
 Best Local Similarly 100.0%; Pred. No. 3.1e-237;
 Matches 900; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtagagcttcaccatattatcttcttcgagatgatacaattctctcggg 60
 Db 1 atgctagagcttcaccatattatcttcttcgagatgatacaattctctcggg 60
 QY 61 atttcacaaatgagcaatctgtagtgagatgcaatctgtagcaatgacagaaga 120
 Db 61 atttcacaaatgagcaatctgtagtgagatgcaatctgtagcaatgacagaaga 120
 QY 121 atggctccgctgagatctctctctctctctctctctctctctctctctctctg 180
 Db 121 atggctccgctgagatctctctctctctctctctctctctctctctctctctg 180
 QY 181 ttctg 240
 Db 181 ttctg 240
 QY 241 aatctgcaatctg 300
 Db 241 aatctgcaatctg 300
 QY 301 gttctg 360
 Db 301 gttctg 360
 QY 361 aggatctacaaagctggtcccaatgagatctctctctctctctctctctctctg 420
 Db 361 aggatctacaaagctggtcccaatgagatctctctctctctctctctctctg 420
 QY 421 tctgtctg 480
 Db 421 tctgtctg 480
 QY 481 tcccaaaaatgcaaatctcaaaaagaagatacactgctacataagattctctctg 540
 Db 481 tcccaaaaatgcaaatctcaaaaagaagatacactgctacataagattctctctg 540
 QY 541 gctgagctg 600
 Db 541 gctgagctg 600
 QY 601 ctgggagagcacaaccggcgaatgagaaacagtgagcggcagcaggtctctctgag 660
 Db 601 ctgggagagcacaaccggcgaatgagaaacagtgagcggcagcaggtctctctgag 660

QY 661 ggtgaccatcagcgtgtgtctctctctctctctctctctctctctctctctccac 720
 Db 661 ggtgaccatcagcgtgtgtctctctctctctctctctctctctctctctctccac 720
 QY 721 tgcataaaagttctctctctctctctctctctctctctctctctctctctg 780
 Db 721 tgcataaaagttctctctctctctctctctctctctctctctctctctctg 780
 QY 781 ttctg 840
 Db 781 ttctg 840
 QY 841 aatcctaattgaaacaaatgcaaaaagttctctctctctctctctctctctg 900
 Db 841 aatcctaattgaaacaaatgcaaaaagttctctctctctctctctctctctg 900

RESULT 2

AAF92528

ID AAF92528 standard; DNA; 1331 BP.

XX

AC AAF92528;

XX

DE 16-MAY-2001 (first entry)

XX

Rat T2R01 nucleotide sequence SEQ ID NO:78.

XX

Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor; taste transduction G-protein coupled receptor; identification; tongue; taste sensory neuron; taste cell; taste modulator; food;

XX

taste signalling pathway; ds.

XX

OS

XX

XX

Rattus sp.

XX

MO200118050-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 1331 BP; 348 A; 289 C; 266 G; 428 T; 0 other;

Query Match 43.4%; Score 390.2; DB 22; Length 1331;
Best Local Similarity 66.1%; Pred. No. 3, 2e-97;
Matches 595; Conservative 0; Mismatches 298; Indels 6; Gaps 2;

```

OY 1 atgctgagctcaccatcattatcttctctctgagctgataaattctcttgg 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 77 atgaggaaggaggaatactctctctctctctctctctctctctctctct 136
OY 61 atttcacaatgacacacattgtgtgtgtgaaatgagctgactgctcagaac 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 137 gctctggcaaatggcctcattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 196
OY 121 atggtcgcgtgactctctctctctctctctctctctctctctctctcag 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 atgcccgcgtgactctctctctctctctctctctctctctctctctcag 256
OY 181 ttcattcttaagtaagtgtatgtatctctctctctctctctctctctc 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 tctatattgttgcaaatgtgtctctctctctctctctctctctctct 316
OY 241 aattgtgcaattctctctctctctctctctctctctctctctctctcag 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 317 aatattactctgtctctctctctctctctctctctctctctctctctc 376
OY 301 gttctctattgtgcaaatgtgtcgcagctgcgcacacacacacacacac 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 377 gttctctattgtgcaaatgtgtcgcagctgcgcacacacacacacacac 436
OY 361 aagatatcaaatgtgtcgcagctgcgcacacacacacacacacacacac 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 437 aggaataccaggtgtgtcgcagctgcgcacacacacacacacacacac 496
OY 421 tgtgtcttccacaagaatgcaaggtgtatgtgtccactctctctctcag 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 497 acttctcattcagagagagagagagagagagagagagagagagagagag 556
OY 481 tcccaaatgacacacacacacacacacacacacacacacacacacacac 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 557 ccttaaaatg---caactcaagtcgaaacagggcagcacaactcagtc 613
OY 541 gctgagctcagctgctctctctctctctctctctctctctctctctct 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 614 ctgtggtctacacacacacacacacacacacacacacacacacacacac 673
OY 601 ctgtggtgagggcagacacacacacacacacacacacacacacacacac 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 674 ctgtggaattatagcagcagatgagacta---tgttaggcacacagga 730
OY 661 ggtgcaaccatcagcagctgtgtctgtctctctctctctctctctctc 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 731 catgtcacaatagtgcaatgtgtctctctctctctctctctctctctc 790
OY 721 tgcataataaaggtttctctctctctctctctctctctctctctctct 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 791 taccagtggtgctgtctctctctctctctctctctctctctctctct 850
OY 781 tctctcattctgtgattgtatataccctctctctctctctctctctct 840
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 851 tctctcttactgtgtattgtatgtataccctctctctctctctctctct 910
OY 841 aatcctaattgaacaaatgcaaaatgctccctctctctctctctctct 899
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 911 aatcctaagctgaacaaatgcaaaatgtctctctctctctctctctct 969

```

RESULT 3
AAF92560
ID AAF92560 standard: DNA; 1698 BP.
XX AAF92560;
AC

```

XX 16-MAY-2001 (first entry)
DT Mouse T2R19 nucleotide sequence SEQ ID NO:142.
XX
DE Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor;
XX taste transduction G-protein coupled receptor; identification; tongue;
XX taste sensory neuron; taste cell; taste modulator; food;
XX taste signalling pathway; ds.
XX
OS Mus sp.
XX
PN W0200118050-A2.
XX
PD 15-MAR-2001.
XX
PF 08-SEP-2000; 2000MO-US24821.
XX
PR 10-SEP-1999; 9905-0393634.
XX
PR 22-FEB-2000; 2000US-0510332.
XX
PA (REGC ) UNIV CALIFORNIA.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zuker CS, Adler JE, Ryba N, Mueller K, Hoon M;
XX WPI; 2001-211396/21.
XX P-PDB; AAB87813.
XX
XX Nucleic acids encoding the T2R family of G-protein coupled taste
XX receptors, useful for identifying taste modulators that can be used in
XX food and pharmaceutical industries to customize taste, for e.g. to
XX decrease the bitter taste of food .
XX
XX Claim 51; Page 231; 249pp; English.
XX
XX AAF92502 to AAF92572 represent nucleic acids which encode taste
XX transduction G-protein coupled receptors designated T2R proteins.
XX AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
XX represent T2R family consensus sequences from the present invention.
XX The T2R proteins are taste modulators. The nucleic acids are useful as
XX probes for the identification of taste cells, as the nucleic acids are
XX specifically expressed in taste cells. They also serve as tools for the
XX generation of taste topographic maps that elucidate the relationship
XX between the taste cells of the tongue and taste sensory neurons leading
XX to taste centres in the brain. The taste modulators are useful for
XX pharmacological and genetic modulation of taste signalling pathways.
XX CC Modulatory compounds comprising T2R proteins can therefore be used in
XX food and pharmaceutical industries to customise taste, for e.g. to
XX decrease the bitter taste of food or drugs.
XX
XX Sequence 1698 BP; 426 A; 397 C; 352 G; 523 T; 0 other;

```

Query Match 43.2%; Score 388.6; DB 22; Length 1698;
Best Local Similarity 66.1%; Pred. No. 9, 7e-97;
Matches 594; Conservative 0; Mismatches 299; Indels 6; Gaps 2;

```

OY 1 atgctgagctcaccatcattatcttctctctgagctgataaattctcttgg 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 515 atgaggaaggacatagctctctctctctctctctctctctctctctct 574
OY 61 atttcacaatgacacacattgtgtgtgtgaaatgagctgactgctcagaac 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 575 gctctggcaaatggcctcattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 634
OY 121 atggtcgcgtgactctctctctctctctctctctctctctctctctcag 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 635 atgcccgcgtgactctctctctctctctctctctctctctctctctcag 694
OY 181 ttcattcttaagtaagtgtatgtatctctctctctctctctctctctcag 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 695 tctatattgttgcaaatgtgtctctctctctctctctctctctctctc 754

```

QY	241	aattgagcgaattctctctatctatctaaatctgaatctgggaactctgctctgcaacatctgctcgc	300
Db	755	aagctacccctctgctccacatctaaacgaacatcgagctctctggcttctggcacacatgctcgtg	814
QY	301	gtctctcatctgctgcacaaagcttgccacgcgtccgctcacccacctctccalcctggcttgaaagatg	360
Db	815	gtctctcatctgctgcacaaagcttgccacgcgtccgctcacccacctctccalcctggcttgaaagatg	874
QY	361	aggaatactccaagactctgtctcccatctgagatgacccctggggctctctctctatagtatctatgatt	420
Db	875	aaggaatactccaagactctgtctcccatctgagatgacccctggggctctctctctatagtatctatgatt	934
QY	421	tgatctctccatctgacgaatactgacgaaggtcttgatggctcccatctcctaagaagaattcttc	480
Db	935	aactctccatctgacgaatactgacgaaggtcttgatggctcccatctcctaagaagaattcttc	994
QY	481	lcccaaaatgcccacaaatctcaaaaagaagaatacactggtctatatacagaatcttctctctgtct	540
Db	995	lctaaaaaat--acaaactcgggtctcagacacgacgcatgccaacatactctccagctctgtctc	1051
QY	541	gctgagctctccagctgacatctgcttatctctctcttctgctctgctctctctgattctctct	600
Db	1052	cttggtgctccacactaccatcttccatctctacactgctctgtctctgtctctctgtctctctcc	1111
QY	601	ctggggaggagaccccccggacaaatcgagaagaacacagtggtccggacgacagggtctccctggacag	660
Db	1112	ctgtggtgaacacacacgcgcgcgcagatgag--gactatggtggtgggaactcagggaacctgacaga	1168
QY	661	ggatgacccacacacgcgcgcgcagctgtctctctatctccctgtccctctgacctccactctccac	720
Db	1169	catgctcctctgctcagctgagatgctctccatctctgctcatctccatccatctccatctccat	1228
QY	721	tgcctgtgtaaaagtcttctg	780
Db	1229	gacatctgctgacgtctctctgacctgctacccaaaggcctccacatcttggaagcagaacctgtgca	1288
QY	781	ttctctacccctctgcatctgatatataccctctctggacacctctccatctcttaattttaa	840
Db	1289	cttcgcttatctgttatctgttatgataccctctcccttaacactcgatctgtcttaattttaa	1348
QY	841	aatctcaaatctgaaacaaatctgacaaaagtctctctctccacacagtaagtgctgtcagctg	899
Db	1349	aacctgaagcttgaaacgaatactgacaaaacgcttccatctgctcatctgtaagtgctgtcagctg	1407
RESULT 4			
ID	AAF92508	standard; DNA; 957 BP.	
AC	AAF92508;		
XX			
DT	16-MAY-2001	(first entry)	
XX			
DE	Human T2R07 nucleotide sequence SEQ ID NO:14.		
KW	Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor; identification; tongue;		
KW	taste transduction G-protein coupled receptor; identification; tongue;		
KW	taste sensory neuron; taste cell; taste modulator; food;		
KW	taste signalling pathway; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200118050-A2.		
XX			
PD	15-MAR-2001.		
XX			
PF	08-SEP-2000; 2000MO-US24821.		
XX			
PR	10-SEP-1999; 99US-0393634.		
PR	22-FEB-2000; 2000US-0510332.		
XX			
PA	(REGC) UNIV CALIFORNIA.		

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PT
 XX Zuker CS, Adler JE, Ryba N, Mueller K, Hoon M;
 XX
 DR MPF; 2001-211396/21.
 XX P-PSDB; AAB87737.
 PT Nucleic acids encoding the T2R family of G-protein coupled taste
 PT receptors, useful for identifying taste modulators that can be used in
 PT food and pharmaceutical industries to customise taste, for e.g. to
 PT decrease the bitter taste of food -
 XX
 PS Disclosure; Page 165-166; 249pp; English.
 XX
 AAF92502 to AAF92572 represent nucleic acids which encode taste
 CC transduction G-protein coupled receptors designated T2R proteins.
 CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
 CC represent T2R family consensus sequences from the present invention.
 CC The T2R proteins are taste modulators. The nucleic acids are useful as
 CC probes for the identification of taste cells, as the nucleic acids are
 CC specifically expressed in taste cells. They also serve as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centres in the brain. The taste modulators are useful for
 CC pharmacological and genetic modulation of taste signalling pathways.
 CC Modulatory compounds comprising T2R proteins can therefore be used in
 CC food and pharmaceutical industries to customise taste, for e.g. to
 CC decrease the bitter taste of food or drugs.
 XX
 Sequence 957 BP; 256 A; 202 C; 192 G; 307 T; 0 other;

[illegible]

to taste centres in the brain. The taste modulators are useful for pharmacological and genetic modulation of taste signalling pathways. Modulatory compounds comprising T2R proteins can therefore be used in food and pharmaceutical industries to customise taste, for e.g. to decrease the bitter taste of food or drugs.

Sequence 951 BP; 218 A; 221 C; 209 G; 303 T; 0 other;

Query Match	13.3%	Score 119.8	DB 22	Length 951
Best Local Similarity	51.3%	Prod. No. 4.5e-23		
Matches 448	Conservative	0	Mismatches 392	Indels 33
				Gaps 6
QY	32	ttctctgcagatgtgtaacaattctctctctgttggaattcttcacaagaatgtgcatctgtgtcgtgtga	91	
DB	32	ttctctgtctggtcactcaatgttcaatctacatctcgtggaaattctgttcaattgtttcaattgtgtgtca	91	
QY	92	atgagcatctatctgtatcctaagcagacagaaaaatggtctcgtctgtgattctctctctctctgtc	151	
DB	92	atgtagtagcagctgtgtctcaaggaacaaagagaatgtctctgtctgtactcatcatcacaacc	151	
QY	152	tggcaggtctctctgaatctctctctcgcagttgttcaactt-----ctacgttaatgtgat	203	
DB	152	tggcagctctctggaatctctctcgtgtgtatattctgactgtatagatttttaataaagaat	211	
QY	204	tgttatctctctcagaagaattcacaatgtgtctgtcgtgaattgtgaattctctt---at	259	
DB	212	tctctccccaacacacacatgtatcttcagggaataataatgcaaaattctatgttctctcgtgacat	271	
QY	260	tctataatgtaatcttgaaacttttggtctgtccacaatggtctcgtgtcttctctatgtgtgccaaag	319	
DB	272	tctcaaaaacacatctgtgcgactttgtgctgtccacactgtctgtgtctctctactatgctcgtcaaaa	331	
QY	320	tgtccagagctctccgttaaccacactctcatatctctgtatgaagaatgaggaataaccagagctgtctc	379	
DB	332	tgtccagagctctctcctcaaccacacactctctctctgtgtctcaagtgaggaattctctaaagggtgat	391	
QY	380	catgtgatagtctctcgtgtgtctgtctgtataatgtatactatgtattgtgttttccataagaat	439	
DB	392	tatgtgaatgctgtctgt	449	
QY	440	atgcagaggtttatgtgttccacatactctcctaagaagaattttctcccaaa-----atgc	491	
DB	450	----tgaagtttaagctctatctctgtctcttaagggaattgagtcacacagaaatgtatgactg	505	
QY	492	ctcgaattcgaaaaaagaagaataacacgtgctatacagaattctctcttctgtgtcgtgaattctc	551	
DB	506	aacacactctcagaagaagaagaagtgaggtatatactgataccatgtctctctggaactctgtgtc	565	
QY	552	agtgcgaattgtctatactctctctctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	607	
DB	566	acctgctctctcttaattgt	625	
QY	608	gtgcacacccgtgcgaatagagaaacacagatgtgctcgtgcagcaggggtctctcgtgcagaggtgtcgc	667	
DB	626	gtgcacacacgtgcagatgt	685	
QY	668	ccatcagcagcgtctgtcgtgtatactcgtgtctctctctctctctctctctctctctctctctctctct	727	
DB	686	aaagaaggtgcacatcagaatacatct	745	
QY	728	taaaagttttctctctctctctctcaaggtttccacacatcgaaggtgtcatctctctctctctca	787	
DB	746	cttctcttaattgtcatcaatttgttataattctctacaaaaaaccaagaatggtctaagaatgtatg	805	
QY	788	tctctgtgattgtgta---tataccctctctgcgaactctctctcatctaatctttagaatac	844	
DB	806	gtggaagtaattgacaatgttttatactctgtcgtgcacacttattctcattctcgtgtggaaca	865	
QY	845	ctaaattgtaaaacaaaatgtcaaaaaagtctctcc	877	
DB	866	gtatagcttgaaagacattgttgaattgtatgtctcc	898	

```

RESUL_T  9
ID AAF92571 standard; DNA: 3775 BP.
XX
XX AAF92571;
XX
XX 16-MAY-2001 (first entry)
XX
XX Mouse T2R30 nucleotide sequence SEQ ID NO:163.
XX
XX Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor;
XX taste transduction G-protein coupled receptor; identification; tongue;
XX taste sensory neuron; taste cell; taste modulator; food;
XX taste signalling pathway; ds.
XX
XX Mus sp.
XX
XX WO200118050-A2.
XX
XX 15-MAR-2001.
XX
XX 08-SEP-2000; 2000WO-US24821.
XX
XX 10-SEP-1999; 99US-039334.
XX PR 22-FEB-2000; 2000US-0510332.
XX
XX (REGC ) UNIV CALIFORNIA.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zuker CS, Adler JE, Ryba N, Mueller K, Hoon M;
XX
XX WPI: 2001-211396/21.
XX DR P-PsDB: AAB87823.
XX
XX Nucleic acids encoding the T2R family of G-protein coupled taste
XX receptors, useful for identifying taste modulators that can be used in
XX food and pharmaceutical industries to customize taste, for e.g. to
XX decrease the bitter taste of food -
XX
XX Claim 51; Page 243-245; 249pp; English.
XX
XX AAF92502 to AAF92572 represent nucleic acids which encode taste
XX transduction G-protein coupled receptors designated T2R proteins.
XX CC AAB87721 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
XX represent T2R family consensus sequences from the present invention.
XX CC The T2R proteins are taste modulators. The nucleic acids are useful as
XX CC probes for the identification of taste cells, as the nucleic acids are
XX CC specifically expressed in taste cells. They also serve as tools for the
XX CC generation of taste topographic maps that elucidate the relationship
XX CC between the taste cells of the tongue and taste sensory neurons leading
XX CC to taste centres in the brain. The taste modulators are useful for
XX CC pharmacological and genetic modulation of taste signalling pathways.
XX CC Modulatory compounds comprising T2R proteins can therefore be used in
XX CC food and pharmaceutical industries to customise taste, for e.g. to
XX CC decrease the bitter taste of food or drugs.
XX
XX SO Sequence 3775 BP; 1312 A; 629 C; 761 G; 1073 T; 0 other;

Query Match 12.1%; Score 108.8; DB 22; Length 3775;
Best Local Similarity 49.0%; Pred. NO. 8.9e-20;
Matches 423; Conservative 0; Mismatches 417; Indels 24; Gaps 4;

19 atatactattctctctctgcaagtgcataacaattctctctctgggaatttcacaagtgcac 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 488 actacacttaactgcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 547
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2Y 79 atgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 138
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 548 atgtgcacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 607
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2Y 139 ctctctctctctgtctgtgcagttctctagaattctctctgtgcagttgttcac-----ttct 190

```


Db	608	atctctccaagttgagccatgctccagaattgtgtctaacagtgtaatactatctagattgt	667
Qy	191	acgttaatgtgattgttattctttcttcatagaattcataatggtcttcgtcaattgtgca	250
Db	668	attatattgtgtcagttatccagacaccttaaaacagaggtaaagaaatgtgagccgtltgac	727
Qy	251	ttctctt----attataaagtatttgaaactttgtctgtccacaaagtgtctggcgttttc	306
Db	728	ttctctcgtgaaactctaccacaaacttaagtgtctgtgtttgttgcacactgctccagcatcttc	787
Qy	307	tattgtgccaaaggtttgcagcgctccgltccaccaactctcatccttggtttgaagaatggagala	366
Db	788	tattattccaagatagaanaaactctctccaccctctttccctctgtatgaataagtgaggaatt	847
Qy	367	tccaagctgtgtcccatgtgatagccctggggtctctctgtatatagtatactatgtattgtgt	426
Db	848	gacaaagcnaattcccaagaactctactgtgacatgtgtatatactctccgtgtttttagccctc	907
Qy	427	ttccatagtcaaatatgtcagaggttatatgtgtcccaactctccaaagaaattttcccca	486
Db	908	ccagtaactgtgaaatctcgtagtattgtattcagaacgttgtgtgtaagcaagagagaata	967
Qy	487	aatgtccacaattccaanaaagaagataacactgtgcatacagaatttctcttgttgcgtgag	546
Db	968	aactctactcttgagatltcaaaagtataaagaagctgtgacatgaccttltcaaggtlaaatctc	1027
Qy	547	ttctcag-----tgcattgctatctctctcttcttggctgtttgtctcttgatttc	597
Db	1028	aacttgttcatgctgttccctcttctctgtctctgtctgtctctcttccctcttgaccttc	1087
Qy	598	tctcttggtgagagcacaccctcgcaaatltagaanaacagltgcccgtgcagaggttccctgcgc	657
Db	1088	ttccctgtgtgagacacaccacagcagatatacaactcagtgtaacaggtgaacaaagatccagcc	1147
Qy	658	agggggtgcacccaatcagcgcggtgtgtcgtatctctgtctctctgtactctactctcc	717
Db	1148	acaacagctctctgtgaagaagcatatgaagcagtaatttctctctctgtccctgttgtgttc	1207
Qy	718	cacgtcagataaagaagtttctctctctctcctaagtattccatactagaaggttcatcctt	777
Db	1208	tactgtccagacttcttccatagcacaactccagctactctttagtcgaagaggtgaattagct	1267
Qy	778	ctgtctctcatccttctgtatggt---ataaccctctgtgacactctctcatcthaatt	834
Db	1268	gtaataatggtgtgacgtatagctctatactatctatccctcaagccattcatattatccctatc	1327
Qy	835	ttaggaatccttaaatgtgaacaa	888
Db	1328	ctgtgggtagtagtaactaaacaa	1351

RESULT	10
AAF92523	
ID	AAF92523 standard; DNA: 960 BP.
XX	
AC	AAF92523:
XX	
DT	16-MAY-2001 (first entry)
XX	
DE	Human T2R30 nucleotide sequence SEQ ID NO:52.
XX	
KW	Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor;
KW	taste transduction G-protein coupled receptor; Identification; tongue;
KW	taste sensory neuron; taste cell; taste modulator; food;
KW	taste signalling pathway; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200118050-A2.
XX	
PD	15-MAR-2001.
XX	

```

FF 08-SEP-2000; 2000OWO-US24821.
PR 10-SEP-1999; 99US-0393634.
PR 22-FEB-2000; 2000OUS-0510352.
XX
PA (REGC ) UNIT CALIFORNIA.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Zuker CS, Adler JE, Ryba N, Mueller K, Hoon M;
XX
XX WPI; 2001-211396/21.
DR P-PSDB: AAB87760.
XX
PT Nucleic acids encoding the T2R family of G-protein coupled taste
PT receptors, useful for identifying taste modulators that can be used in
PT food and pharmaceutical industries to customize taste, for e.g. to
PT decrease the bitter taste of food .
XX
PS Disclosure; Page 180-181; 249pp; English.
XX
XX AAF92502 to AAF92572 represent nucleic acids which encode taste
CC transduction G-protein coupled receptors designated T2R proteins.
CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
CC represent T2R family consensus sequences from the present invention.
CC The T2R proteins are taste modulators. The nucleic acids are useful as
CC probes for the identification of taste cells, as the nucleic acids are
CC specifically expressed in taste cells. They also serve as tools for the
CC generation of taste topographic maps that elucidate the relationship
CC between the taste cells of the tongue and taste sensory neurons leading
CC to taste centres in the brain. The taste modulators are useful for
CC pharmacological and genetic modulation of taste signalling pathways.
CC Modulatory compounds comprising T2R proteins can therefore be used in
CC food and pharmaceutical industries to customize taste, for e.g. to
CC decrease the bitter taste of food or drugs.
XX
SQ Sequence 960 BP; 256 A; 197 C; 183 G; 324 T; 0 other;
XX
Query Match 12.1%; Score 108.6; DB 22; Length 960;
Best Local Similarity 49.4%; Pred. No. 5.4e-20;
Matches 431; Conservative 0; Mismatches 424; Indels 18; Gaps 5;
OY 1 atgctagagctcaccatcattatcattctctcttcagtgatcaaatctctctgg 60
DB 1 atgataactttctcaccatccattttctccatctcgtgagtcgttaccattgtcttga 60
OY 61 atttcacaatgagatcatatggtggttgatgagtcattgactgtatcaagaacagaana 120
DB 61 aattttcccaatggtcaltatagctctagtaaatcattcatttgatggtggtcaagaacagaag 120
OY 121 atggtcgcgtagctctctctctctctctgctgcgagttccagattcttcagattcttcgagctg 180
DB 121 atctctcagctgacccaatctctacgtcctcgtggtggtctccaggttggtttactctgg 180
OY 181 ttcaactt-----ctacgttaagtgatgttattcttcttcattatagaattcatalg 231
DB 181 gtcatattatcatgtgtatgcaaatggtttaatcgaacttataatagttcagaagta 240
OY 232 tgttcgcgaattgagcaattctctatttataaattgaaattggaacttggctgcaca 291
DB 232 agactctgtgctctcaaatctcaagaacaatacaacacalltaagaatctggtcttact 300
OY 292 tggcgcgcgatttctatatgtgcagaagttgcgcagcgcgcgcacacactcttacttgcg 351
DB 301 agccctcagcatatttatttctcgaagattgcgaattctccacacttaattttctccac 360
OY 352 ttgaagatgagatataccaagcttggtcccatgagatcctcgtggggtctctgataatga 411
DB 361 ttaagaagagagatttagagtggttctcgtgagatacgtttgggtcccttggtattttg 420
OY 412 tctatgatttggtttctccatagcaaatatgcaggttatagtgtcccatacttctaagg 471
DB 421 atttgttaattctgctgtatataacccttgatg-accgttggttgacaagaatagaag 479

```


Db 1663 gtaaccacgcttgcaactaataaccccttactctgtccctgttatcttttctctta 1722
 Qy 592 attctctcttgaggagcacaccggcaaatggaacaaagtgccgcgcaagaggt 651
 Db 1723 atctgtctcttgtaaacatctcaagaagatgcagttcgaacagatcctccagat 1782
 Qy 652 cctggcaggggtgcaaccatcagcgctgtctctatctctgtcctccgatctctac 711
 Db 1783 tccaaccacgaagtcacataaagcttgcgaacggtgacccctctctctgttattt 1842
 Qy 712 -----ttccccaatgcatgataaagtttctctctctcctcaagttccacatcaga 765
 Db 1843 gctgttactctctgtccccaatcacaatcgatlttggaggaagcgtgtaaac 1902
 Qy 766 aggttcaactctctgtctctcattcctgtgtatataaccctctgcgaactcttc 825
 Db 1903 gaacctgtccatcagtcagcaaacactacatcgaattatataccctcattcattc 1962
 Qy 826 atcttaatttaggaatccttaattgaacaaatgcaaaaaagttccctccacagat 885
 Db 1963 atcccaatttgagggaagcaagaagctgaacagaccccttcttgcatttgtgtcagatt 2022
 Qy 886 aagtgctg 893
 Db 2023 aagtgctg 2030

RESULT 12
 AAF92515
 ID AAF92515 standard: DNA; 903 BP.
 AC AAF92515:
 DT 16-MAY-2001 (first entry)
 DE Human T2R15 nucleotide sequence SEQ ID NO:29.
 XX
 XX
 KW Human: rat; mouse; T2R; taste receptor; G-protein coupled receptor;
 KW taste transduction G-protein coupled receptor; identification; tongue;
 KW taste sensory neuron; taste cell; taste modulator; food;
 KW taste signalling pathway; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200118050-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 08-SEP-2000; 2000MO-US24821.
 XX
 PR 10-SEP-1999; 99US-0393634.
 PR 22-FEB-2000; 2000US-0510332.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Zuker CS, Adler JE, Ryba N, Mueller K, Hoon M;
 DR MPI: 2001-211396/21.
 DR P-PSDB: AAB87745.
 XX
 PT Nucleic acids encoding the T2R family of G-protein coupled taste
 PT receptors, useful for identifying taste modulators that can be used in
 PT food and pharmaceutical industries to customize taste, for e.g. to
 PT decrease the bitter taste of food -
 XX
 PS Disclosure: Page 172-173; 249pp; English.
 CC AAF92502 to AAF92572 represent nucleic acids which encode taste
 CC transduction G-protein coupled receptors designated T2R proteins.
 CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830
 CC represent T2R family consensus sequences from the present invention.
 CC The T2R proteins are taste modulators. The nucleic acids are useful as

CC probes for the identification of taste cells, as the nucleic acids are
 CC specifically expressed in taste cells. They also serve as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centres in the brain. The taste modulators are useful for
 CC pharmacological and genetic modulation of taste signalling pathways.
 CC Modulatory compounds comprising T2R proteins can therefore be used in
 CC food and pharmaceutical industries to customise taste, for e.g. to
 CC decrease the bitter taste of food or drugs.
 XX
 SQ Sequence 903 BP; 237 A; 178 C; 171 G; 317 T; 0 other;

Query Match 11.9%; Score 107.2; DB 22; Length 903;
 Best local Similarity 48.9%; Pred. No. 1.3e-19;
 Matches 444; Conservative 0; Mismatches 443; Indels 21; Gaps 5;

Qy 1 atctagaagctcacatcattatcttctcttgcagtgatacaattctcttgg 60
 Db 1 atgataactttctacccaatcatttctccattagtagtgattatgttcttgg 60
 Qy 61 atttcacaatgcatcattgtgtgtgtaatgcatgactgatacagaacagaa 120
 Db 61 aatttgcataatggcttcaatgtagtggtaaatcattgagtgagtcagaagaaag 120
 Qy 121 atggctcgcgtgatactctcttctctgtctgcaggttcataaatcttctcagtg 180
 Db 121 atctccttgcagcaaatctcactgcctgcagctgccacagagtgttctctg 180
 Qy 181 ttcatcttcaagtt-----aatgtagtgtatctcttcatagaattcatalgt 232
 Db 181 gtaattattacattgtgtagcaacgttttgtaaccagttgataagtttagagta 240
 Qy 233 gtctcggaattgagcaattcctctatttaaatgtaattggaac-cttgctgcaca 291
 Db 241 agaattcactatataatgcctggctgtaaccacattcagcatcgtgtgctact 300
 Qy 292 tggctcgagtttctatctgtgcacaggtgcagctccgcacacattcattcgtg 351
 Db 301 agctcagcatatttatttccccaagattgcacattccccaatttattcttcac 360
 Qy 352 ttgaagatgaggatacccaagcgtgtcccatgtagtacctggtggtctctgata 411
 Db 361 ttaaaaagggaattaaagatgltcattccagttagtactatgtggtcttgtatt 420
 Qy 412 tctatgttgtgttcttcataagaaatagcaaggtttatggtccatacttctaag 471
 Db 421 gttgtcatcttctgtgtgtaaacatgagtg-agaatagtggaacaaagata 479
 Qy 472 aaatttctcccaaaatgcacacattcaaaaagaagatacactggtcatagatttc 531
 Db 480 aacgtgagttggtgagtaaatg-----agtatccgaacaccccttcagatagact 534
 Qy 532 tcttctgtcgtgagttcagtgcaatgcttactcttctcttctgttctgtctg 591
 Db 535 gtaaccacgcttgcaaatcattaccccttactctgtccctgttactcttctctta 594
 Qy 592 atttctctctggaggagcacaccggcaaatggaacacagcgtgcgagagaggtc 651
 Db 595 atctgtctcttggtgtaaacatcctcaagaagatgagttcattgcaagagatc 654
 Qy 652 cctggcaggggtgcaaccacagcgctgtctctatctcctgtcctcctgactctac 711
 Db 655 tccaaccacgaagtcacataaagcttgcgaacggtgacccctcctctgttattt 714
 Qy 712 -----ttccccaatgcatgataaagtttctctctcctcaagttcacaatcaga 765
 Db 715 gctgttactctctgtccccaatcacaatcgatlttggaggaagcgtgtaaac 774
 Qy 766 aggttcaactctctgtctcattcctctgtgtatataccctctgcgaactcttc 825
 Db 775 gaacctgtccatcagtcagcaaacactacatgcaattatatacccttattcattc 834

CC specifically expressed in taste cells. They also serve as tools for the
CC generation of taste topographic maps that elucidate the relationship
CC between the taste cells of the tongue and taste sensory neurons leading
CC to taste centres in the brain. The taste modulators are useful for
CC pharmacological and genetic modulation of taste signalling pathways.
CC Modulatory compounds comprising T2R proteins can therefore be used in
CC food and pharmaceutical industries to customise taste, for e.g. to
CC decrease the bitter taste of food or drugs.

XX Sequence 1316 BP; 384 A; 226 C; 259 G; 447 T; 0 other;

Query Match 11.3%; Score 101.4; DB 22; Length 1316;
Best Local Similarity 47.5%; Pred. No. 5.9e-18;
Matches 411; Conservative 0; Mismatches 436; Indels 18; Gaps 3;

QY 47 aattctctcttgggaatttcacaaatgcatatctgtgtggaatggaatgcatgactga 106
DB 225 aattctaaatgcatctcttggaatggaatggaatggaatggaatggaatggaatgga 284
QY -107 tcaagcaacgaagaatggtccgctggaatctctctctctctctctctctctctctctctct 166
DB 285 tcaaggaaggaaggaatct 344
QY 167 tctctctgagctgtctcatct 226
DB 345 tgcgaactgtctgttgaatataatgtaagtgtgtggaatattgtctctctctctctctctct 404
QY 227 tcaatgtctctggaatctgtgcaatctctctctctctctctctctctctctctctctctct 274
DB 405 ggaatgcatgatagaagatctgaatataatggaatggaatggaatggaatggaatggaatgga 464
QY 275 aacttggtctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctg 334
DB 465 gttctctgtctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctg 524
QY 335 acccaactctctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctg 394
DB 525 acccaactctctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctg 584
QY 395 ggtctctgtctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctg 454
DB 585 tgcctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaatggt 644
QY 455 tcccaactctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctg 514
DB 645 tcaatggaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 704
QY 515 tggcctaaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 574
DB 705 tgcctctggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaatggt 761
QY 575 tggcctaaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 634
DB 762 tcaatggaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 821
QY 635 tggcctaaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 694
DB 822 tcaatggaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 881
QY 695 tggcctaaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 754
DB 882 tcaatggaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 941
QY 755 tggcctaaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 811
DB 942 tcaatggaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 1001
QY 812 tggcctaaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 871
DB 1002 tcaatggaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 1061
QY 872 tggcctaaatggaatggtctggaatggtctggaatggtctggaatggtctggaatggtctggaat 896

DB 1062 tctccctggtggtgaggtgagca 1086

Search completed: December 11, 2001, 02:39:56
Job time: 5768 sec.

Tue Dec 11 09:54:24 2001

us-09-510-332-2.rng

THIS PAGE BLANK (USPTO)

Gencore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2001, 22:44:08 ; Search time 1381.63 Seconds

(without alignments)
6999.849 Million cell updates/sec

Title: US-09-510-332-2

Perfect score: 900
Sequence: 1 atgcagtagctcaccatc.....acagtaagtcgtcagatga 900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estrov:*
9: em_hlc:*
10: gq_estl:*
11: gq_estl2:*
12: gq_hlc:*
13: gq_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pin:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750.4	83.4	806	11	BG209561 RST9087
2	681	75.7	742	13	AQ308694 CTRBI-EI-
3	649	72.1	650	13	AQ316999 CTRBI-EI-
4	514.2	57.1	758	11	BG203110 RST22479
5	293	32.6	310	13	AQ277039 CTRBI-EI-
6	249.4	27.7	261	11	BG192000 RST11106
7	248.6	24.8	321	11	AA416581 AA416581
8	223.2	8.9	554	10	AL037695 AL037695
9	80.2	6.8	383	13	AQ103952 HS_5107_A
10	61.2	6.0	589	13	AQ590563 HS_5400_A
11	53.8	5.9	787	13	AZ739176 RPCI-24-1
12	52.8				

C 13	48.4	5.4	1101	13	CNS00KX5	AL078382 Drosophila
C 14	47.6	5.3	399	11	N52978	N52978 yv30B01.S1
C 15	47	5.2	652	13	BH045040	BH045040 RPCI-24-3
C 16	46.6	5.2	399	11	H81901	H81901 yu78a04.s1
C 17	46.6	5.2	528	13	AQ719085	AQ719085 HS_5514_B
C 18	45.4	5.0	949	13	CNS04A1H	AL281906 Tetradon
C 19	45.2	5.0	878	13	CNS0187R	AL108993 Drosophila
C 20	44.4	4.9	656	13	A2416308	A2416308 1M0191D04
C 21	42.8	4.8	928	13	CNS00DKY	AL071865 Drosophila
C 22	42.6	4.7	639	13	CNS0170D	AL108367 Drosophila
C 23	42.4	4.7	752	13	CNS0477G	AL127621 Tetradon
C 24	42.2	4.7	679	13	CNS0164V	AL106297 Drosophila
C 25	42.2	4.7	889	13	A2531644	A2531644 ENTB5591R
C 26	42.2	4.7	1034	13	CNS06M08	AL05798 T3 end of
C 27	42	4.7	849	13	A2671816	A2671816 ENTMR06TR
C 28	41.8	4.6	459	10	AA187386	AA187386 zp69C02.r
C 29	41.8	4.6	943	11	BF530656	BF530656 602071951
C 30	41.6	4.6	997	13	CNS005TE	AL060767 Drosophila
C 31	41.6	4.6	780	10	BE642964	BE642964 Cr12.7.10
C 32	41.6	4.6	891	13	A2671751	A2671751 ENTME49TF
C 33	41.6	4.6	901	13	A2531629	A2531629 ENTME49TF
C 34	41.6	4.6	928	13	CNS01119	AL099687 Drosophila
C 35	41.4	4.6	541	13	CNS04KUS	AL295309 Tetradon
C 36	41.4	4.6	809	13	CNS00A7H	AL054820 Drosophila
C 37	41.4	4.6	858	13	CNS010K6	AL099072 Drosophila
C 38	41.4	4.6	1004	13	CNS005PG	AL060500 Drosophila
C 39	41.2	4.6	518	13	A2374617	A2374617 1M0127F01
C 40	41.2	4.6	628	13	BH095829	BH095829 RPCI-24-2
C 41	41.2	4.6	784	10	AL525973	AL525973 AL525973
C 42	41.2	4.6	851	13	CNS009FC	AL053504 Drosophila
C 43	41.2	4.6	1092	13	CNS020K7	AL175696 Tetradon
C 44	41	4.6	224	10	BH130880	BH130880
C 45	41	4.6	498	13	AQ616191	AQ616191 HS_5148_A

ALIGNMENTS

RESULT 1
BG209561
LOCUS
DEFINITION RST9087 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG209561
VERSION BG209561.1 GI:13731248
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 806)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Dahl, T., Thornton, M., Ramachandran, R., Whittington, J., Lerner, L., Krishoc, D., McElligott, K., Clark, S., Mays, R., Smith, E., Veloso, N., Hess, J., Cottrill, K., Lo, K., Offenbacher, J., Danzig, J., and Ducar, M.
TITLE Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression
JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) In press
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scaine@atersys.com
High quality sequence stop: 557.
FEATURES
source
location/Qualifiers
1..806
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression',

Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT 172 a 182 c 160 g 284 t 8 others
ORIGIN

Query Match 83.4% Score 750.4; DB 11; Length 806;
Best Local Similarity 96.5% Pred. No. 9.7e-152;
Matches 763; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

45 acaattctctctggagatcttcacaaatgacatctgtggtggtgaatggcatctact 104
14 ANNTTCTCTTGTGATTTTCACAAATGGCATATGTGGTGGATGGCATTTGACTT 73
105 galcaagacagaaaaatggtctcgctggatctcctctctctctgctggcagttctag 164
74 GATCAAGCACAGAAAAATGCCCTCGATCTCTTCTTCTTCTGCGAGTTTCAG 133
165 aattctctgagctgttcaatctctctctctctctctctctctctctctctctct 224
134 AATTTTCTGCGATGTTGATCTTACGTTAATGATGTTATTTATCTTCTCATGAAATT 193
225 calcatgtctctgcaatctgcaatctctctctctctctctctctctctctctctct 284
194 CATCATGTTGTTGCGAATGTGCAATCTCTTATTTAATGAAATGGAACTTTGGCT 253
285 hgcacatggtctcggtctctctctctctctctctctctctctctctctctctct 344
254 TCCACATGCGCTCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 313
345 catctggttgaagaagaagataaccaagctggccacatgacatcctggagctctct 404
314 CATCTGTTGATGAAGATGAGATATCCAACTGCTCCCATGAGATCTCTGGGTCTGT 373
405 atagatctatgattctgtctctctctctctctctctctctctctctctctctct 464
374 AATGTATCTATGATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 433
465 cctaaaggaattctctcccaaaatgacacaaatcacaagaagaagaacatgctatca 524
434 CCTAAGCAAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 493
525 gattct 584
494 GATTTCTCTTTTGTGCTGAGATTCAGTTCATGCTTCTTCTTCTTCTTCTTCTT 553
585 gctcttgaattct 644
554 GCTCTTGAATTTACTACTGAGGAGGACACCCGCGAATGAGAAACACAGTGGCGGCA 613
645 cagggtctctgcaaggggtgacacccatcagcgtctgctctctctctctctctct 704
614 CAGGCTTCTGCGAGGAGGACACCCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
705 cct 764
674 CCTCTACTCTCTCCCATGATGATANAAGTTTCTCTCTCTTCTTCTTCTTCTTCT 733
765 aaggttcaatct 824
734 AAGGGTCATCTCTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 793
825 catcttaattt 835
794 CTTCCTAATTTT 804

RESULT 2
LOCUS A0308694 742 bp DNA GSS 22-DEC-1998
DEFINITION CITBI-E1-253086.1F CITBI-E1 Homo sapiens genomic clone 253086, DNA

sequence.
ACCESSION A0308694 GI:4040728
VERSION A0308694.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 742)
Adams, M.D., Rounsley, S.D., Zhao, S., Baas, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.

TITLE
JOURNAL
COMMENT
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@ligr.org

Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..742
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="253086"
/clone_1id="CITBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

BASE COUNT 176 a 175 c 144 g 247 t
ORIGIN

Query Match 75.7% Score 681; DB 13; Length 742;
Best Local Similarity 100.0% Pred. No. 8.7e-137;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

220 gaattcattatgctgtctgcaatgctgaatctctctctctctctctctctctctct 279
2 GAATTCATCATGTTGTTGCGAATGTGCAATTCCTTATTTATTAATGAATTTGAACTT 61
280 tggcttgcacatggtctcggtctctctctctctctctctctctctctctctctctct 339
62 TGGCTTGCACATGCTGCGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 121
340 cctctcatctgttgaagaatgagatatccaaagctggtcccatgagatgacctggggtct 399
122 CTCTTCATCTGTTGAGATGAGATATCCAAAGCTGTCCTCATGATGATCTCTGGGCTT 181
400 ctgcataatgataatgattctgtctctctctctctctctctctctctctctctctct 459
182 CTGCTATATGATGATGATTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 241
460 taattctcfaagaaattctctcccaaaatgacacaaatcacaagaagaatacaatgct 519
242 TACTTCTTAAGAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 301
520 atcagaattctctctgt 579
302 ATACAGATTTTCTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 361
580 gtttgcctctgattct 639
362 GTTTTGTCTTGTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 421

QY 640 ggcagcagggttcctgagcagggtgacccatcagcgctgtctatctccttc 699
| | | | |
Db 422 ggcagcagggttcctgagcagggtgacccatcagcgctgtctatctccttc 481
| | | | |
QY 700 ctgacccctactctcccaactgacatgaatgcttctctctctctctctctctc 759
| | | | |
Db 482 ctgacccctactctcccaactgacatgaatgcttctctctctctctctctc 541
| | | | |
QY 760 atcagaaggttcaactctctctctctctctctctctctctctctctctctc 819
| | | | |
Db 542 atcagaaggttcaactctctctctctctctctctctctctctctctctctc 601
| | | | |
QY 820 tctctcctctaatcttgaggaatctcctaatgacacaaatgcaaaagtctctc 879
| | | | |
Db 602 tctctcctctaatcttgaggaatctcctaatgacacaaatgcaaaagtctctc 661
| | | | |
QY 880 cacagtaagtgctgacagcga 900
| | | | |
Db 662 cacagtaagtgctgacagcga 682
| | | | |
RESULT 3
LOCUS A0316999 650 bp DNA GSS 22-DEC-1998
DEFINITION CITRBI-E1-2530B7.TF CITRBI-E1 Homo sapiens genomic clone 2530B7, DNA
sequence.
ACCESSION A0316999
VERSION A0316999.1 GI:4040265
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 650)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeel@igf.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21.
Class: BAC ends.
FEATURES
Source Location/Qualifiers
1. 650
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2530B7"
/clone_lib="CITRBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT 149 a 152 c 125 g 224 t
ORIGIN
Query Match 72.1%; Score 649; DB 13; Length 650;
Best Local Similarity 100.0%; Pred. No. 6.9e-130;
Matches 649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 220 gaattcatatgtgtctgcaatgtgcaatctcttattataaagaatggaactt 279
| | | | |
Db 2 gaattcatatgtgtctgcaatgtgcaatctcttattataaagaatggaactt 61
| | | | |

QY 280 tggcttgcacatagctcgcggcttctctatgtgcgaagttgccagcgtccgaccca 339
| | | | |
Db 62 tggcttgcacatagctcgcggcttctctatgtgcgaagttgccagcgtccgaccca 121
| | | | |
QY 340 ctctcatctgttgaagatgagatataccagctgtgtccatgagatgagatcctgggtc 399
| | | | |
Db 122 ctctcatctgttgaagatgagatataccagctgtgtccatgagatgagatcctgggtc 181
| | | | |
QY 400 ctgctataatgataatgattgtgtgttccatagcaaatatgcaaggttattgtccca 459
| | | | |
Db 182 ctgctataatgataatgattgtgtgttccatagcaaatatgcaaggttattgtccca 241
| | | | |
QY 460 tacttctaaaggaaatttctcccaaatgccaatctcaaaatgagaatgacatgct 519
| | | | |
Db 242 tacttctaaaggaaatttctcccaaatgccaatctcaaaatgagaatgacatgct 301
| | | | |
QY 520 atacagaattctctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 579
| | | | |
Db 302 atacagaattctctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 361
| | | | |
QY 580 gtttgcctctgatttctctctctgaggagacaccccgcaaatgagaacagtgcc 639
| | | | |
Db 362 gtttgcctctgatttctctctctgaggagacaccccgcaaatgagaacagtgcc 421
| | | | |
QY 640 ggcagcagggttcctgagcagggtgacccatcagcgctgtctatctccttc 699
| | | | |
Db 422 ggcagcagggttcctgagcagggtgacccatcagcgctgtctatctccttc 481
| | | | |
QY 700 ctgacccctactctcccaactgacatgaatgcttctctctctctctctctc 759
| | | | |
Db 482 ctgacccctactctcccaactgacatgaatgcttctctctctctctctctc 541
| | | | |
QY 760 atcagaaggttcaactctctctctctctctctctctctctctctctctctc 819
| | | | |
Db 542 atcagaaggttcaactctctctctctctctctctctctctctctctctctc 601
| | | | |
QY 820 tctctcctctaatcttgaggaatctcctaatgacacaaatgcaaaagtctctc 868
| | | | |
Db 602 tctctcctctaatcttgaggaatctcctaatgacacaaatgcaaaagtctctc 650
| | | | |
RESULT 4
LOCUS BG203110 758 bp mRNA EST 21-APR-2001
DEFINITION RST22479 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG203110
VERSION BG203110.1 GI:13724797
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 758)
Hartington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Dahl,T., Thornton,M., Ramachandran,R., Whittington,J.,
Lerner,L., Krashoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,
Velooso,N., Hess,J., Cochren,K., Lo,K., Offenbacher,J., Danzig,O.
and Ducar,M.
TITLE Creation of Genome-wide Protein Expression Libraries using Random
Activation of Gene Expression
JOURNAL Nat. Biotechnol. 19 (5), 440 (2001) In press
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave. Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: sca@atersys.com
High quality sequence stop: 440.
FEATURES
Source Location/Qualifiers
1. 758
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

/clone.lib-"Athersys RAGE Library"
/cell_line="H1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is H1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in H1080 under normal circumstances."
BASE COUNT
ORIGIN
164 a 169 c 165 g 240 t

```

Query Match 57.1%;	Score 514.2;	DB 11;	Length 758;
Best Local Similarity 95.8%;	Pred. No. 7.3e-101;		
Matches 528;	Conservative 0;	Mismatches 23;	Indels 0;
			Gaps 0

Accession	Sequence	Position
QY	gaagagatcaaatctctctcttggaattccacaanaaygacatctggtgtygaatggc	96
Db	gcagcagatcatttcttcttgsgattttcacaaatggcatctttggatggatggc	265
QY	. 97 attgacttgatcaagacacagaaaalbgctcgcgtgatactctctctctctgga	156
Db	atttgacttgatcagacacagaaaalggcgcgcgtgattctcttcttcttggca	325
QY	gtctctagaatcttctcgaattgttcacatctctcaagttaaagtgatgttatctcttc	226
Db	gtttctcagcaatttttttctgcagattgtttcattcttctacggttaattctgattctttac	385

397	fcctctcctatataatctatctatctaatcttctatcttccctacacgaatactacgaacgctctatctatc	456	
0y			
Db	386	atagaaatttatctatctgtgtgtcttgccgaatttgcaatttctcttattctaatatgaatttgga	445
0y	277	ctttggtcttgcacatgagctcggcgtcttctctatcttgcgaaggttgcagcgtctcac	336
Db	446	cttttggtcttgcacatgagcctgcggcgttttctatattgagccaaagtttgcaacgcttgcgcgac	505
0y	337	ccactctcctacatctgttgtgaagatgagagatatccaaagcttggtcccatgatatgatactccgggg	396
Db	506	ccactctttttatctgctttggaagatgagagattttccaccctgtgcccctaatggaatgattcctgggg	565

Db	566	TCCTCTGCTATATGATCTATGATTTGTGTCTCCATAGCAAAATATGCAGGAGATTTATGGCC	625
Qy	457	cccaacttccttaaggaaattttctccccaanaatgccaacaaatcaaaaaagaataacacig	516
Db	626	CCATACCTTTCTAAGGACATTTCTCTCCCAAAATGCCCCCATTTTAAAAAGAAAGATACACTG	685
Qy	517	gctatacagaatttcctccttttgtctgcbaagttctcagttgccaatlgcttalcctccctttt	576
Db	686	CCCATACACATTTTTCCTCCTATGCTTGCCTGAAATACACGCCCAATGCCTTATCTGCTTTT	745
Qy	577	gcctgttttgct	587
Db	746	GCCTGATGCT	756

RESULT	5		
LOCUS	AQ277039		
DEFINITION	AQ277039 310 bp DNA	GSS	22-NOV-1998
	CITR1-El-2516A7.TF CITR1-El	Homo sapiens genomic clone 2516A7,	DNA
	sequence.		

ACCESSION	AQ277039	GI:3903235
VERSION	AQ277039.1	
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS

1 (bases 1 to 310)

Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K., Berry, K., Gaenger, D., Suh, E., Wible, C., Shizuya, H., Simon, M., and Venter, J.C.

TITLE	Use of a random human BAC End Sequence Database for sequence-Ready
JOURNAL	Map Building
COMMENT	Unpublished (1998)
	Other_GSSS: CITRBT-EI-2516A7_TIR

Contact: Mark Rodenas
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mrodenas@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/hungen/Bac_end_search/Bac_end_search.html
Seq primer: M13-21
Class: BAC ends.

FEATURES	SOURCE	Location/Qualifiers
	1.	310
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="2516A7"	
	/clone_1ib="C17B1-E1"	
	/sex="male"	
	/cell_type="sperm"	
	/note="type: pBel0AC11; site_1: EcoRI; site_2: EcoRI; CalTech Human BAC Library D"	
BASE COUNT	75 a	77 c 69 g 88 t 1 others
ORIGIN		

Query Match	32.6%	Score 293;	DB 13;	Length 310;
Best Local Similarity	96.5%	Pred. No. 3,5e-53;		
Matches 299;	Conservative	0;	Mismatches 11;	Indels 0;
				Gaps 0;

QY 368 ccaagctctggtcccatggatgataccctgggtctctgctataatgatactatgatgttgltt 427
Db 1 ccaagctctggtcccatggatgataccctgggtctctgctataatgatactatgatgttgatc 60
QY 428 tccatagaagaataatgacaggtttatggttcccaatacttcccaaggaanatttctcccaaa 487
Db 61 tccatagaagaataatgacaggtttatggttcccaatacttcccaaggaanatttctcccaaa 120
QY 488 atgcacacattccaagaagaatacactctggtctatacagaattctctcttggtctgaatc 547
Db 121 atgcacacattccaagaagaagaatacactctggtctatacagaattctctcttggtctgaatc 180
QY 548 tctcagtgccattgcttatactcttccttcttgctgtttctgcttgatcttcctctcgggaa 607
Db 181 tctcagtgccattgcttatactcttccttcttgctgtttctgcttgatcttcctctcgggaa 240
QY 608 ggcacaccccggaataatgagaacaacagtgccggcagcaggttctctgacaggtgtgcac 667
Db 241 ggcacaccccggaataatgagaacaacagtgccggcagcaggttctctgacaggtgtgcac 300
QY 668 ccatacagcgc 677
Db 301 ccatacagcgc 310

RESULT	6
BG192000	
LOCUS	BG192000 261 bp mRNA EST 21-APR-2001

DEFINITION:	KTSL1106 Altermys KAGE LIDBAY	HOMO	SEPIENS	CDSN	miRNA sequence:
ACCESSION:	BC192000				
VERSION:	BC192000.1	GI:	13713687		
KEYWORDS:	EST				
SOURCE:	human				

ORGANISMS
invertebrates: Echinodermata; Cephalopoda; Mollusca; Crustacea; Insecta; Arachnida; Chelicerata; Onychophora; Eutherozoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Eutherozoa; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
1 (bases 1 to 261)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Dahl, T., Thornton, M., Ramchandran, R., Whittington, J.,

Veloso, N., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, J., and Ducar, M.
 Creation of Genome-wide Protein Expression Libraries using Random
 Activation of Gene Expression
 Nat. Biotechnol. 19 (3), 440 (2001) In press
 Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atersys.com

FEATURES
 source
 Location/Qualifiers
 1. 321
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Athersys RAGE Library"
 /cell_line="HT1080"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

BASE COUNT 96 a 59 c 54 g 112 t
 ORIGIN
 Query Match 24.8% Score 223.2; DB 11; Length 321;
 Best Local Similarity 98.3%; Pred. No. 3.8e-38;
 Matches 236; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 42 gacaaatctctctctggatcttcacaaatgacatctggtggtggaatgacatga 101
 |||||
 Db 6 CATCAATTTCTTCTTGTA-TTTCAGCAATGCGATCTTGTGGCGATGCGATGA 64

OY 102 ctgacacagacagaaatggtcgcgtgcatccctctctctctgctgcaattc 161
 |||||
 Db 65 CTGATTAAGACAGAAATGCGTCCGCTGGATCTCTCTCTCTGCGCAGTTTC 124

OY 162 tagaatctctctgacgtgtgtcatctctacgttaatgagatggtatctctcataga 221
 |||||
 Db 125 TACGAATTTTCTGCGATGTCATCTTACGTTAATGTGATGTTATCTTTCATAGA 184

OY 222 attcaccatgctgctgcaatctgcaatctctctctctctctctgcaattg 281
 |||||
 Db 185 ATTCACTATGCTTCTGCGATGTCATCTTCTTATTATTAATGATGGAATTTG 244

RESULT 9
 LOCUS AL037695 554 bp mRNA EST 29-FEB-2000
 DEFINITION DKFZ56400672_r1 564 (synonym: hfb2) Homo sapiens cDNA clone
 DKFZ56400672.5, mRNA sequence.
 ACCESSION AL037695
 VERSION AL037695.3 GI:5928268
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 TITLE EST (Blum, et al.)
 JOURNAL Unpublished (1999)
 COMMENT On Jul 7, 1999 this sequence version replaced gi:5865546.
 CONTACT: Blum H

MIPS
 Am Klopferpitz 18a D-8152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by LMU (Ludwig Maximilians University,

Munich/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 s1 sequence also available.
 This clone (DKFZ56400672) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 Location/Qualifiers
 1. 554
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZ56400672"
 /clone_lib="564 (synonym: hfb2)"
 /tissue_type="brain"
 /dev_stage="fetal"
 /lab_host="X1-2blue"
 /note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

BASE COUNT 154 a 96 c 108 g 196 t
 ORIGIN
 Query Match 8.9% Score 80.2; DB 10; Length 554;
 Best Local Similarity 48.9%; Pred. No. 2.2e-07;
 Matches 250; Conservative 0; Mismatches 253; Indels 8; Gaps 1;

OY 22 atcattctctctgagtgatatacaattctctctggtatcttcacaaatgcatcat 81
 |||||
 Db 5 ATTTATCAATTTCTGATGTTGATTTGTTCTTGGAATGTTGCCAATGCTCATCA 64

OY 82 gtgtgtgtaatgagcatctgactgatacaagacagaaatggtcgcgtgcatctct 141
 |||||
 Db 65 GCTGTAGTAATGTCAGTAAGACTGGTTAAGACACAAAGATCTCTCACTGACCAAT 124

OY 142 ctctctctgcaagctctcagaat-----cttctgcaattctcactctacg 193
 |||||
 Db 125 GTCACTGCTCTGCGATCTTCCAGAAATGTTTACTTGTATATTTATACATTGATG 184

OY 194 ttaatgtaattgtatactctctcagaatcatcatcagtgctcgcgaattgcaattc 253
 |||||
 Db 185 CAAGTGTATTAATTAAGCTTTATATATATTAAGAAATGTAATGTTCTCTATGCT 244

OY 254 tctattataaagaattggaactctgctgcacacagtcgcgtctctcattgtg 313
 |||||
 Db 245 CGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 304

OY 314 ccaaggtctcagcgtccgcacacacacacacacacacacacacacacacacacac 373
 |||||
 Db 305 TCAGATTTGCCAATTTCTCAATTTATTTATTTCTCCACCTTAAGAGAAATTAAGAG 364

OY 374 tggccatgatactcctcgtggtctctgctatagatgatactgattctcacta 433
 |||||
 Db 365 TTCTTCTTGTATCTGTTGGGCTCTTGTATTTTATTTGTATTTCTCTGTGATGA 424

OY 434 gcaaatatgaggtgtatgtgtcccatctctcagaagaaattctctcccaaatgcca 493
 |||||
 Db 425 CCAATGGATGACAGGTGCTGTGACAAAAGAAATTTGAAGCAATGTGACTTGGAGGATC 484

OY 494 caattcaaaagaatatactgactatata 524
 |||||
 Db 485 GAATGAGGAATGCAATACACCTTCAACACA 515

RESULT 10
 LOCUS AQ103952 383 bp DNA GSS 28-AUG-1998
 DEFINITION HS_3107_AL_C05_T7 CIT Approved Human Genomic Sperm Library D Homo
 sapiens genomic clone Plate-3107 COL-9 Row-E, DNA sequence.
 ACCESSION AQ103952
 VERSION AQ103952.1 GI:3478888
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 383)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence tagged Connector
 Plate: 3107 row: E column: 9
 Class: BAC ends
 High quality sequence stop: 383.

FEATURES
 source 1. 383
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate=3107 Col=9 Row=E"
 /clone_1ib="CIT Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="Organ: sperm; Vector: pBelobAC11; BAC clones in
 E-Coli DH10B"
 BASE COUNT 112 a 69 c 71 g 128 t 3 others
 ORIGIN

Query Match 6.8%; Score 61.2; DB 13; Length 383;
 Best Local Similarity 57.4%; Pred. No. 0.0029;
 Matches 108; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
 Oy 1 atgctagagctcaccatcatatcttctcttcgagtgatcaaatctcttcg 60
 Db 184 ATGATMACTTTTCCCATCATTTTCCATTTCAATAGTGTATATTGTTATTTGGG 243
 Oy 61 atttcaaatgagatcatgtgtgtggaagcattgactgataagcacaga 120
 Db 244 AAATTTGCTTAATGGCTTATGATGTGTAATTCATTTGAGTGGTCAAGACAAAG 303
 Oy 121 atggctccgctgagatctctctctctctgctgagcttctagaattcttcgagttg 180
 Db 304 ATCTCCTTGTGACCAATTTCTCAGTCTCTGCGCTCTCAGAGTNTGCTCTCTCG 363
 Oy 181 ttcactct 188
 Db 364 GTGGTATT 371

RESULT 11
 LOCUS AO590563 589 bp DNA GSS 08-JUN-1999
 DEFINITION HS_5400_A1_G04_T7A RPCI-11 Human Male BAC library Homo sapiens
 genomic clone Plate=976 Col=7 Row=W, DNA sequence.
 ACCESSION AO590563
 VERSION AO590563.1 GI:5022215
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 589)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/Ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end web server:
<http://www.htsc.washington.edu>
 Plate: 976 row: W column: 7
 Seq primer: 77
 Class: BAC ends
 High quality sequence stop: 589.

FEATURES
 source 1. 589
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate=976 Col=7 Row=W"
 /clone_1ib="RPCI-11 Human Male BAC library"
 /sex="male"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"
 BASE COUNT 166 a 98 c 117 g 199 t 9 others
 ORIGIN

Query Match 6.0%; Score 53.8; DB 13; Length 589;
 Best Local Similarity 60.7%; Pred. No. 0.11;
 Matches 102; Conservative 0; Mismatches 65; Indels 1; Gaps 1;
 Oy 1 atgctagagctcaccatcatatcttctcttcgagtgatcaaatctcttcg 60
 Db 311 ATGATGCTTTTCTGCTCATTCATTCATCAATTCGATGAGTTCG-ATTGTCTTGA 369
 Oy 61 atttcaaatgagatcatgtgtgtggaagcattgactgataagcacaga 120
 Db 370 AATGTGCCAATGGCTTCAATAGCTAGTAATGTCATTGACTGTTAAACACAGAAAG 429
 Oy 121 atggctccgctgagatctctctctctctgctgagcttctagaattcttcgagttg 168
 Db 430 ATCTCCTCAGCTGACCAATTTCTCAGTCTCTGCTGCTCCAGAAAT 477

RESULT 12
 LOCUS A2739176 787 bp DNA GSS 25-JAN-2001
 DEFINITION RPCI-24-154D22.TV RPCI-24 Mus musculus genomic clone RPCI-24-154D22
 , DNA sequence.
 ACCESSION A2739176
 VERSION A2739176.1 GI:12509031
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 787)
 AUTHORS Zhao, S., Niemman, W., Malek, J., Shatsman, S., Akiret, B., Levins, M.,
 Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E.,
 Russell, D., de Jong, P. and Fraser, C.M.
 TITLE Mouse BAC End Sequences from Library RPCI-24
 JOURNAL unpublished (1999)
 COMMENT Other-GSSs: RPCI-24-154D22.TJ
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA

Tue Dec 11 09:54:25 2001

us-09-510-332-2.ts